Test Patterns Compression Technique Based on a Dedicated SAT-based ATPG

Jiří Balcárek
Dept. of Computer Science & Engineering
Czech Technical University in Prague, FEL
Prague, Czech Republic
balcaji2@fel.cvut.cz

Petr Fišer, Jan Schmidt
Faculty of Information Technology
Czech Technical University in Prague
Prague, Czech Republic
fiserp@fit.cvut.cz, schmidt@fit.cvut.cz

Abstract—In this paper we propose a new method of test patterns compression based on a design of a dedicated SAT-based ATPG (Automatic Test Pattern Generator). This compression method is targeted to systems on chip (SoCs) provided with the P1500 test standard. The RESPIN architecture can be used for test patterns decomposition. These are decompressed in the scan chain during the test, no additional hardware is required. The main idea is based on finding the best overlap of test patterns. During the test generation, we are trying to efficiently generate patterns as candidates for an overlap, unlike other methods, which are based on efficient overlapping of pre-generated test patterns. The proposed algorithm takes advantage of an implicit test representation as a SAT problem instance. We show results of test patterns compression obtained for standard ISCAS’85 and ’89 benchmark circuits. The results are compared with competitive test compression methods.

Keywords: Test patterns compression, testing, ATPG, SAT

I. INTRODUCTION

Testing of digital circuits is quite a difficult task, for a huge amount test data needed to be delivered to the circuit under test (CUT). With the growing complexity of designs, scan-based test techniques are becoming a standard. The test patterns are shifted into the chain of scan registers (scan chain) through a serial interface and the circuit under test response is shifted out to the response compactor.

Generally, we can use a deterministic or pseudorandom test patterns set for testing. Both have their advantages and drawbacks. Pseudorandom testing may be easily realized on the chip by a linear-feedback shift register (LFSR) [1, 2] or by other automata [3]. Test patterns generated in such a way cover most of easily-detectable faults, but the patterns can be quite inefficient in covering random pattern resistant faults. Also a great number of test patterns are needed to generate. It means that the time consumption may grow up significantly.

On the other hand, we can generate deterministic test patterns for all detectable faults. These patterns, having a considerable data volume, need to be sent into scan chains. This data transfer is realized by a TAM (Test Access Mechanism), which creates an interface between ATE (Automatic Test Equipment) and the on-chip test mechanism.

Design requirements force us to make the TAM as narrow as possible, but sending test patterns through a narrow TAM may cause a considerable growth of the testing time. That is why the compression of the test patterns is needed to decrease the bandwidth between ATE and TAM.

There are several methods used for a test patterns compression. Many of these methods are based on using of some encoding, such as statistical codes [2, 4, 5, 6, 7], run-length codes [4, 5, 6, 8, 9], and Golomb codes [10], others are based on XOR networks [11, 12], hybrid patterns [13], EDT (Embedded Deterministic Test) [14] and reuse of scan chains [15]. Another approaches are based on test patterns compaction [16, 17] and overlapping [18, 19, 20, 21].

In this paper we introduce a novel test patterns compression algorithm SAT-Compress based on a design of a dedicated SAT-ATPG [22, 23]. The compression algorithm exploits an overlapping of test patterns. Unlike other compression methods based on this principle [18, 19, 20, 21], we do not use test patterns pre-generated by an ATPG; we generate them on the fly to reach the locally best overlap and maximize the compression. Test patterns compressed this way can be easily decompressed by the RESPIN decompression architecture [24], which is intended to test system on chip (SoC) cores by reusing scan chains.

The paper is organized as follows: first a brief summary of the previous work on test patterns compression is presented in Section II. Then the representation of the test vector set as an instance of the SAT problem is considered and basic principles of SAT-based ATPGs [22] are described (see Section III). The RESPIN decompression architecture is shown in Section IV. A new test patterns compression algorithm based on dedicated SAT-ATPG is introduced in Section V, and finally the experimental results and comparison with other compression methods are shown (see Section VI).

II. PREVIOUS WORK

The proposed test patterns compression method is basically based on finding the best overlap of test patterns pre-generated by an ATPG (Fig. 1). The test patterns are serially shifted into the scan chain. This idea was described in [21] for the first time. This algorithm generally tries to find contiguous and consecutive test patterns having the maximum overlap. Deterministic test patterns are generated by an ATPG and compacted. Patterns in the scan chain are checked whether they match with one or more test patterns which were not employed in the sequence yet. In [20], the pattern overlapping problem is
converted into a Traveling Salesman Problem (TSP), for which different heuristics have been proposed.

<table>
<thead>
<tr>
<th>Test patterns</th>
<th>Overlapped test patterns</th>
</tr>
</thead>
<tbody>
<tr>
<td>0010--</td>
<td>0010--</td>
</tr>
<tr>
<td>010--</td>
<td>010--</td>
</tr>
<tr>
<td>101--</td>
<td>101--</td>
</tr>
<tr>
<td>01010</td>
<td>01010</td>
</tr>
<tr>
<td>100--</td>
<td>100--</td>
</tr>
<tr>
<td>01111</td>
<td>01111</td>
</tr>
<tr>
<td>11100</td>
<td>11100</td>
</tr>
<tr>
<td>10001</td>
<td>10001</td>
</tr>
</tbody>
</table>

Test bitstream: 0010010101010001
Compressed test bitstream: 0010101001110001

The COMPAS [18] test patterns compression tool is based on a similar approach, but it does not use compacted test patterns. Test patterns that are to be compressed are pre-generated by an ATPG as well. However, these test patterns should contain as many don’t care bits (DC bits, means not specified) as possible. For this purpose, one test pattern for each fault from the fault list is generated. Greater number of don’t care bits grants the algorithm much more possibilities to combine test patterns and reach better compression. Other improvements are the simulation after every test pattern application and searching for best successors of a given starting pattern (usually an all-zero pattern). These improvements make COMPAS very efficient in comparison with other compression tools, see. Table II. The only weakness of COMPAS is the need for don’t care bits. The number of DCs and the fact that the algorithm fully relies on a pre-generated test set can also affect the test compression ratio. When the test patterns are highly specified (they contain only few don’t care’s), it is much harder for COMPAS to find a good overlap of the test patterns and the efficiency of the test patterns encoding decreases which cause greater memory consumption [25].

Another compression technique based on the RESPIN architecture is presented in [15]. Suitable test patterns are produced by a standard ATPG tool performing dynamic compaction, while constraints to circuit inputs are applied. In contrast to this, we apply these constraints to the SAT instances, in a SAT-based ATPG.

The state-of-the-art compression/decompression architecture used in industry is the Embedded Deterministic Test (EDT) [14]. Here a high test compression is achieved by employing a dedicated but generic test decompressor. The compressed test patterns serve as seeds for a pseudo-random pattern generator (“ring register”), where they are decompressed and further distributed to scan-chains using a XOR-network structure (“phase-shifter”). The compressed test patterns are obtained as a solution of a set of linear equations. Similarly to the previously mentioned test compression methods, the test patterns to be compressed need to be precomputed by an ATPG. Again, high amount of test don’t cares is essential for achieving a good compression ratio [14].

### III. SAT-BASED ATPG AND ITS PRINCIPLES

#### A. Circuit to SAT Conversion

Conversion of a digital circuit to a CNF (Conjunctive Normal Form) is an essential task for our SAT-based compression algorithm. The SAT instance in CNF represents in our case the whole set of test patterns detecting a given fault. Each SAT instance is described as a Boolean formula in CNF. A CNF \( \Phi \) with \( m \) Boolean variables is a conjunction of \( n \) clauses, where each clause is a disjunction of literals. Each literal is a Boolean variable or its complement. The CNF \( \Phi \) is satisfiable, if there exists some assignment of variables, for which all clauses are satisfied.

A Boolean variable is assigned to each signal in the circuit \( C \). For each gate, the CNF \( \Phi_g \) is derived from its characteristic function. The CNF \( \Phi_{g} \) representing the circuit’s function is constructed as a conjunction of the CNFs of all gates \( g_1, \ldots, g_n \in C \):

\[
\Phi_C = \prod_{i=1}^{n} \Phi_{g_i}
\]

To generate the test for a fault \( F \), the characteristic function \( \Phi_f \) of the faulty circuit is generated. Only a necessary part of the circuit has to be included in the CNF \( \Phi_f \), for reasons of efficiency; circuit parts not propagating the fault may be omitted. Finally the SAT instance for a fault \( F \) is obtained as a product of \( \Phi_C, \Phi_f \) and a characteristic function of a XOR of the fault-free and faulty circuit outputs:

\[
\Phi_{test}^F = \Phi_C \cdot \Phi_f \cdot \Phi_{XOR}
\]

The \( \Phi_{test}^F \) solutions represent the whole set of test patterns detecting the fault \( F \). If \( \Phi_{test}^F \) is unsatisfiable, the fault \( F \) is undetectable. More information can be found in [22, 23].

Example 1. A circuit having three gates is shown in Fig. 2. To detect a given fault, two copies of the circuit are created: the fault-free circuit and the circuit with a particular fault. In order to detect the fault, output values of these two circuits must differ. This is indicated by XORing their outputs (\( X \) and \( X' \)). Characteristic functions are derived from logic functions of the gates. For example, let us consider the AND gate \( D=A\land B \). For any two functions \( P \) and \( Q \), \( P=Q \) is equivalent to \((P\Rightarrow Q)\land(Q\Rightarrow P)\). In this way the AND gate characteristic function is constructed as \((D\Rightarrow (A\land B))\land((A\land B)\Rightarrow D)\). Next we transform this expression to CNF, obtaining \((\neg D\lor A)\land(\neg D\lor B)\lor(D\lor A\lor B)\).
The decompression of compressed test patterns is conducted by the RESPIN (REusing Scan chains for test Pattern decompression) [15, 24] architecture. This architecture is intended for testing systems on chip (SoCs). Basically, scan chains of different cores are reused for updating of the content of the tested core scan chain (SC). RESPIN is based on the IEEE 1500 proposed standard for testing of embedded cores [26]. Each core has a serial test access, which is a mandatory element of IEEE 1500 and a parallel access. A core test wrapper, which is the interface between the embedded core (core terminals) and its system on chip environment (rest of the integrated circuit and TAM), is defined. By definition, the core test wrapper is implemented on-chip and should have three mandatory modes:

1) Normal operation – the core is connected to its system IC (integrated circuit) environment and the wrapper is transparent.

2) Core-internal test – the TAM is connected to the core such that test patterns can be applied at the cores inputs and responses are observed at cores outputs.

3) Core-external test - the TAM is connected to the interconnect wiring and logic such that a test patterns can be applied at cores outputs and responses are observed at the cores inputs.

An example of RESPIN architecture for one SC is shown in Fig. 3.

This architecture consists of ETC (Embedded Test Core) and CUT. During the test, both are isolated from the rest of the SoC by a test wrapper. The decompression of test patterns is performed in the ETC, which is supposed to be a reused scan chain of another core (ETC SC). This ETC SC is extended only by a single multiplexer (controlled by signal test - t_c) and a feedback wire. Reusing of scan chains grants a minimal decompression hardware overhead.

The CUT scan chain is the scan chain of the tested core. It is filled by the decompressed test patterns from the ETC SC and, after application of the functional clock, the responses are shifted out to the signature analyzer (SA). SA compares responses with the expected signature and the circuit is marked as fault-free or faulty.

The control signal t_c (test) of the ETC SC is also connected to the scan enable signal of the CUT SC. This control signal selects between the serial mode (t_c = 0) and the circular mode (t_c = 1) of the ETC SC and between capture mode (t_c = 0) and scan mode (t_c = 1) of the CUT SC.

Table I. shows that while the CUT SC captures the system response (t_c = 0, capture mode) and shifts it out to the signature analyzer (t_c = 1, scan mode). The ETC SC is in serial mode (t_c = 0) and loads one new bit from the encoded test data (compressed test bitstream). Thus only one bit per test pattern is scanned into the ETC SC from the ATE. Whenever the CUT SC scans in test data (t_c = 1, scan mode) from the ETC SC, the ETC SC is in circular mode and the feedback is used to shift valid part of the test pattern back to the ETC SC and reuse it in new test pattern. For using cores as ETC SC both their shift and capture clocks must be compatible with the CUT SC clocks and ETC SC must have at least as many scan cells as the CUT SC.

More specific details of the RESPIN architecture can be found in [15, 24].
signature analyzer (SA) to be evaluated. At the same time a new test pattern is shifted from ETC SC into the CUT SC. Moreover, the hardware overhead is minimal since the decompression is made by reusing of the SCs of non-tested cores and no additional decompression hardware is required.

<table>
<thead>
<tr>
<th>t = scan enable</th>
<th>0</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>CUT</td>
<td>capture</td>
<td>scan mode</td>
</tr>
<tr>
<td>ETC</td>
<td>serial</td>
<td>circular</td>
</tr>
</tbody>
</table>

**TABLE I.** OPERATION MODES OF ETC SC AND CUT SC [15, 24]

![Testing Architecture Diagram](image)

**Figure 3.** RESPIN testing architecture [15, 24]

V. THE SAT-COMPRESS ALGORITHM PRINCIPLES

A. Basic Idea

In our new approach we try to eliminate weaknesses of COMPAS. The main idea is not to overlap test patterns pre-generated by an ATPG, but to generate most suitable test patterns on the fly, to reach the (locally) best overlap. The basic question is how to find these test patterns. Each fault has its set of test patterns by which it is detected. If we were able to pick the right pattern for each fault in the right order, we could have reached the best possible compression of the test patterns.

Because explicit computation and storing of all these test patterns is inefficient (and mostly even infeasible), we were forced to find another, more efficient way of test patterns set representation.

We have researched possibilities of implicit representations of test patterns. We have found that we can take advantage of principles of SAT-based ATPGs and efficiently represent all test patterns for one fault *implicitly*, by one SAT problem instance in a CNF. The CNF test set representation is much less memory consuming than a standard tabular test set representation. Our proposed test set compression algorithm sequentially generates the compressed test bitstream. A SAT instance in CNF is generated for each fault. This SAT instance contains both variables representing the circuit’s primary inputs and variables representing internal signals. The test pattern is determined by values of the primary input variables in the SAT solution; values of internal variables are of no significance. The variables not present in the generated CNF or in the SAT solution are set as don’t cares (DCs). A SAT instance is solved with constraints given by the test pattern generated in the preceding algorithm step. A satisfiable solution of a SAT instance forms the next suitable test pattern for an overlap. Higher time consumption can be expected, because this method requires a repeated SAT solving. Our aim, however, is to find the best overlap and maximize the test patterns compression, as testing time, not test generation time, forms a bottleneck.

B. The SAT-COMPRESS Algorithm Description

We try to find the best overlap by gradually building the compressed test patterns bitstream for the scan chain. The basic algorithm is shown in Fig. 5.

1) First, we generate a complete fault list (FL).

2) A zero state (all-zero test pattern) or the test pattern covering any fault from the fault list is used as the initial test pattern.

3) The pattern is simulated and all detected faults are deleted from the fault list. The leftmost bit of the pattern goes to the resulting bitstream (Fig. 4).

4) The pattern is shifted left and a DC bit is put to its rightmost position. This is the mask for the next pattern (Fig. 4, next pattern mask). The mask constraints the values of primary inputs (PI) in subsequent SAT instances. Only care bits generate the PI constraints.

5) To generate the next test pattern having the highest overlap with the previously generated one, a CNF for each fault in the fault list is generated, while the primary input variables are set according the mask. The orders of faults are of no significance. If a CNF for a fault is satisfiable for a given assignment of primary variables, a new test pattern is obtained. If none of these CNFs is satisfiable, the pattern is shifted one more bit left, which generates a new mask of SAT constraints.

6) These operations (3-6) are performed while the fault list is not empty or until all care bits from the mask of primary inputs setting are shifted out while there is still no satisfiable CNF (rest of the fault in FL is undetectable).

![Next pattern generation example](image)

**Figure 4.** Next pattern generation example
1. Generate FL
   \((FL = \text{fault list})\)
2. \(TP = 0\)
   \((TP = \text{test pattern} = \text{all-zero seed})\)
3. FL = FL - detected_by_simulation(TP)
   \(\text{mask} = \text{DC}\)
   \((\text{DC} = \text{don't care})\)
4. mask = TP[1..n-1] + DC
5. do
   for each fault in FL {
     Create CNF
     Apply mask to PIs in CNF
     if CNF is SAT {
       break the for loop
     }
   }
   \(TP = \text{CNF Solution}\)
   FL = FL - detected_by_simulation(TP)
   compressed_bitstream += TP[0]
   mask = TP[1..n-1] + DC
6. while (FL[i]=0 or (mask==DC and All_CNFS==UNSAT))

Figure 5. The SAT-Compress algorithm

VI. EXPERIMENTAL RESULTS

The experimental results and a comparison of our algorithm with similar compression methods are presented in the following two subsections. Unfortunately, comparison with [15] could not be performed, since only results for 20 parallel scan-chains are presented in [15].

The measurement was performed on Intel Xenon CPU - 2GHz with 4GB RAM.

A. Comparison of Different Compression Techniques

A comparison of SAT-Compress with other state-of-the-art test compression techniques is presented in Table II. The first column “Bench.” represents the benchmark name. The compressed test lengths in bits, for seven different competitive methods, are shown then. A comparison of only seven biggest ISCAS’89 circuits is shown, since no more relevant data was available to us. The last column shows the compressed test data size for our proposed compression tool SAT-Compress. An all-zero initial test pattern for both COMPAS and SAT-Compress is used, thus the results are not influenced by different initial states.

It can be concluded from Table II., that our proposed algorithm can reach better compression of test patterns than most of the presented techniques; we are able to obtain results such as the state-of-the-art compression methods EDT or COMPAS and in some cases even better. The time consumptions of SAT-Compress are shown in Table III. The comparison between presented tools was not possible to measure because of the unavailability of source codes.

B. Comparison COMPAS and SAT-Compress Algorithms

In this Subsection we will present a more detailed comparison of SAT-Compress and COMPAS, as a representative of test compression algorithms based on overlapping of patterns.

The first column “Bench.” presents the name of the benchmark. The compressed test lengths generated by COMPAS and SAT-Compress, both starting with an all-zero test pattern, are shown in the column “bits”. Then we have tried to repeatedly run the algorithms starting with different

Results for ISCAS’85 and ’89 benchmarks [27, 28] are presented in Table III.
initial test patterns. The average compressed test patterns lengths are shown in the columns “avg.” and average variations of compressed test patterns lengths are shown in the columns “var.”. The time consumptions of test patterns compression for SAT-Compress algorithm in seconds are shown in the “time” column. The results of this measurement show that the bitstream length for both tools significantly depends on the initial test pattern and starting with an all-zero seed (which is the default setting for COMPAS) can produce outstandingly poor results, out of the range of the variation (see, e.g., s298 for COMPAS). In some cases, our tool reaches much better compression than COMPAS, but it may also fail. For an unknown reason the efficiency of proposed algorithm is much better for ISCAS’85 benchmarks than for ISCAS’89 ones. As can be seen from Table III., the time consumption may be considerable for larger circuits, but we suppose, that scalability of proposed algorithm may grow up by using of more scan chains and division of the circuit into smaller parts.

Figs. 6, 7, and 8 show examples of distributions of bitstream lengths using different starting patterns for COMPAS and SAT-Compress. In both cases we always start with a test pattern covering one particular fault. The number of restarts is equal to the number of faults.

As can be seen, the compressed bitstream length seems to have Gaussian-like distribution and the only difference between SAT-Compress and COMPAS characteristics is their displacement. This Gaussian distribution of the compressed bitstream lengths for different starting seeds is a common characteristic of these two tools for all tested benchmarks. It can be seen that the selection of the initial test pattern has a crucial impact on the resulting compressed test length, for both algorithms. The way of its proper choice will be a topic of our further investigation.

VII. CONCLUSION

New test patterns compression algorithm (SAT-Compress) based on a modification of SAT-based ATPG is presented. This algorithm utilizes a CNF (Conjunctive Normal Form) implicit representation of test patterns and tries to compress the test patterns by overlapping. In contrast to competitive state-of-the-art test compression techniques, the proposed algorithm does not rely on a pre-generated test set; most suitable test patterns are being generated on the fly.

The test decompression is based on a generic RESPIN architecture, where the patterns are being decompressed in the tested circuit scan-chain. Thus, the circuit needs not be modified to apply the test.

The proposed method was compared with seven state-of-the-art test compression algorithms and a detailed comparison with COMPAS has been made. The comparison results seem to be promising – SAT-Compress achieved the best test compression ratio for many benchmark circuits. Moreover, there are yet many ways of possible improvements. These will be a topic of our further research. Possibilities of application of the method to multiple scan-chain designs will be also studied in the future.

ACKNOWLEDGMENT

This research has been supported by MSMT under research program MSM6840770014 and by the grant of the Czech Grant Agency GA102/09/1668.
REFERENCES


---

**TABLE II. COMPARISON OF THE TEST DATA AMOUNT FOR DIFFERENT COMPRESSION TECHNIQUES**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>s5378</td>
<td>20,758</td>
<td>5,147</td>
<td>6,180</td>
<td>14,572</td>
<td>12,346</td>
<td>-</td>
<td>17,332</td>
<td>2,148</td>
<td>2,407</td>
</tr>
<tr>
<td>s9234</td>
<td>25,935</td>
<td>19,912</td>
<td>12,112</td>
<td>27,111</td>
<td>22,152</td>
<td>-</td>
<td>17,198</td>
<td>11,594</td>
<td>9,928</td>
</tr>
<tr>
<td>s13207</td>
<td>163,100</td>
<td>52,741</td>
<td>11,285</td>
<td>109,772</td>
<td>30,880</td>
<td>10,585</td>
<td>26,004</td>
<td>4,163</td>
<td>10,457</td>
</tr>
<tr>
<td>s15850</td>
<td>58,656</td>
<td>49,163</td>
<td>12,438</td>
<td>32,758</td>
<td>26,000</td>
<td>9,805</td>
<td>32,226</td>
<td>8,234</td>
<td>12,987</td>
</tr>
<tr>
<td>s35932</td>
<td>21,156</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>22,744</td>
<td>-</td>
<td>-</td>
<td>1,860</td>
<td>5,096</td>
</tr>
<tr>
<td>s38417</td>
<td>113,152</td>
<td>172,216</td>
<td>34,767</td>
<td>96,269</td>
<td>93,466</td>
<td>31,458</td>
<td>89,132</td>
<td>24,198</td>
<td>19,291</td>
</tr>
<tr>
<td>s38584</td>
<td>161,040</td>
<td>128,046</td>
<td>29,397</td>
<td>96,056</td>
<td>77,812</td>
<td>18,568</td>
<td>63,232</td>
<td>7,291</td>
<td>14,271</td>
</tr>
</tbody>
</table>